

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-029774
 (43)Date of publication of application : 31.01.2003

(51)Int.Cl. G10L 13/04
 G06F 13/00
 G10L 13/00
 G10L 13/06
 G10L 19/00

(21)Application number : 2001-219513

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 19.07.2001

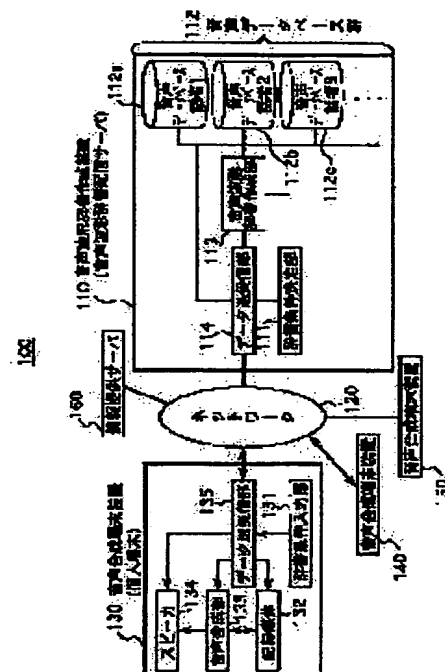
(72)Inventor : MOCHIZUKI AKIRA

(54) VOICE WAVEFORM DICTIONARY DISTRIBUTION SYSTEM, VOICE WAVEFORM DICTIONARY PREPARING DEVICE, AND VOICE SYNTHESIZING TERMINAL EQUIPMENT

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a voice waveform dictionary distribution system, a voice waveform dictionary preparing device and voice synthesizing terminal equipment, by which a text information such as electronic mail can be read in the tone color of a speaker requested by a user.

SOLUTION: In the voice waveform dictionary distribution system 100, when the user of voice synthesizing terminal equipment 130 inputs the preparation conditions of a requested voice waveform dictionary from a dictionary condition input part 131, in a voice waveform dictionary preparing device 110, the voice waveform dictionary requested by the user of the voice synthesizing terminal equipment 130 is prepared by using a plurality of voice databases such as voice databases 112a-112c, and the prepared voice waveform dictionary is recorded through a network 120 onto a recording medium 132 of the voice synthesizing terminal equipment 130.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-29774

(P2003-29774A)

(43) 公開日 平成15年1月31日 (2003.1.31)

(51) Int.Cl.	識別記号	F I	ターミナル (参考)
G 1 0 L 13/04		G 0 6 F 13/00	6 0 5 Q 5 D 0 4 5
G 0 6 F 13/00	6 0 5	G 1 0 L 5/02	G
G 1 0 L 13/00		3/00	E
13/06		5/04	E
19/00		9/18	J
審査請求 未請求 請求項の数13 O L (全 22 頁) 最終頁に続く			

(21) 出願番号 特願2001-219513(P2001-219513)

(22) 出願日 平成13年7月19日 (2001.7.19)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 望月 亮

神奈川県横浜市港北区綱島東四丁目3番1

号 松下通信工業株式会社内

(74) 代理人 100072604

弁理士 有我 軍一郎

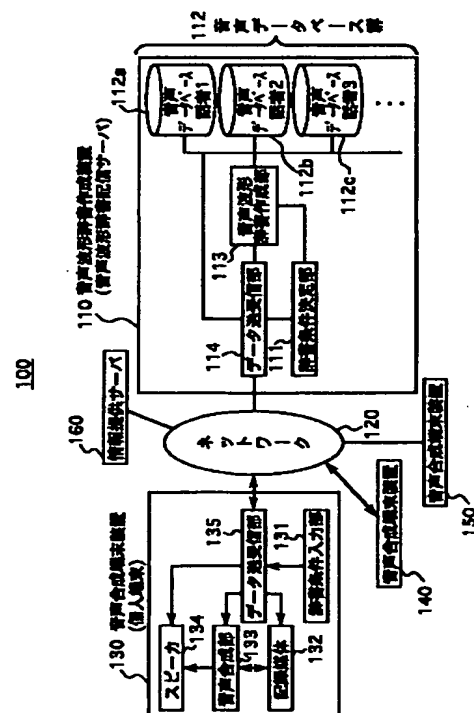
Fターム(参考) 5D045 AB26

(54) 【発明の名称】 音声波形辞書配信システム、音声波形辞書作成装置、及び音声合成端末装置

(57) 【要約】

【課題】 利用者の要求する話者の音色で電子メールなどのテキスト情報を読み上げることが可能な音声波形辞書配信システム、音声波形辞書作成装置、及び音声合成端末装置を提供すること。

【解決手段】 音声波形辞書配信システム100は、音声合成端末装置130の利用者が、要求する音声波形辞書の作成条件を辞書条件入力部131から入力すると、音声波形辞書作成装置110において、音声データベース112a~112cなどの複数の音声データベースを用いて、音声合成端末装置130の利用者が要求する音声波形辞書を作成し、作成した音声波形辞書をネットワーク120を介して音声合成端末装置130の記録媒体132に記録する。



【特許請求の範囲】

【請求項 1】 音声合成に使用される音声波形辞書を作成する音声波形辞書作成装置と、前記音声波形辞書をネットワーク経由で収得して音声合成を行う音声合成端末装置とから構成され、

前記音声波形辞書作成装置は、複数の音声データベースと、前記音声合成端末装置から受信する前記音声波形辞書に対する要求条件に基づいて、前記音声波形辞書の作成条件を決定する辞書条件決定手段と、前記音声データベースの中から前記作成条件に基づいた音声データベースを用いて前記音声波形辞書を作成する音声波形辞書作成手段と、前記ネットワークを介して、前記要求条件を受信し、前記音声波形辞書を送信する条件受信辞書送信手段とを備え、

前記音声合成端末装置は、前記要求条件を入力する辞書条件入力手段と、前記ネットワークを介して、前記要求条件を送信し、前記音声波形辞書を受信する条件送信辞書受信手段と、前記音声波形辞書を記録する音声波形辞書記録手段と、前記音声波形辞書記録手段に記録された前記音声波形辞書を用いて音声合成を行う音声合成手段と、前記音声合成手段で音声合成した音声信号を出力する音声出力手段とを備えることを特徴とする音声波形辞書配信システム。

【請求項 2】 前記音声波形辞書作成装置は、前記音声波形辞書作成手段で作成した前記音声波形辞書を用いて音声合成を実行する作成辞書音声合成手段を備え、前記音声合成端末装置は、前記ネットワークを介して、前記作成辞書音声合成手段で音声合成した音声信号を受信し、受信した前記音声信号を前記音声出力手段で出力することを特徴とする請求項 1 に記載の音声波形辞書配信システム。

【請求項 3】 複数の音声データベースと、音声合成に使用される音声波形辞書に対する要求条件に基づいて、前記音声波形辞書の作成条件を決定する辞書条件決定手段と、前記音声データベースの中から前記作成条件に基づいた音声データベースを用いて前記音声波形辞書を作成する音声波形辞書作成手段と、ネットワークを介して、前記要求条件を受信し、前記音声波形辞書を送信する条件受信辞書送信手段とを備えたことを特徴とする音声波形辞書作成装置。

【請求項 4】 前記音声波形辞書作成手段は、前記音声データベースに予め付属させられた付属情報を参照して音声波形を選出する音声波形選出手段と、前記音声波形選出手段によって選出された前記音声波形を所定の形式に変換するパラメータ化手段と、前記音声データベースから前記音声波形に付属させられた前記付属情報を抽出する音声波形付属情報抽出手段と、前記パラメータ化手段でパラメータ化された前記音声波形、及び、前記音声波形付属情報抽出手段によって抽出された前記付属情報を用いて、前記音声波形辞書を構築する辞書構築手段と

を備えることを特徴とする請求項 3 に記載の音声波形辞書作成装置。

【請求項 5】 前記パラメータ化手段は、前記音声波形を音源特性と声道特性とに分離して、それぞれを独立にパラメータ化することを特徴とする請求項 4 に記載の音声波形辞書作成装置。

【請求項 6】 前記パラメータ化手段は、前記音声波形をピッチ波形単位に分割して抽出するピッチ波形抽出手段と、前記ピッチ波形抽出手段によって抽出された前記ピッチ波形に対して、類似度が所定値以上の前記ピッチ波形同士を分類するピッチ波形分類手段と、前記ピッチ波形分類手段によって分類された前記ピッチ波形の中から前記音声波形辞書に代表として登録する代表ピッチ波形を作成する代表ピッチ波形作成手段とを備えることを特徴とする請求項 4 に記載の音声波形辞書作成装置。

【請求項 7】 前記音声波形選出手段は、予め音声データに付属させられた音素境界ラベル情報を参照して合成単位の境界ラベルを付与し、前記合成単位を作成する合成単位作成手段と、前記合成単位作成手段で作成された前記合成単位の前記音声波形を前記音声波形辞書に登録する候補として抽出する候補波形抽出手段と、前記候補波形抽出手段で抽出された前記音声波形の中から前記音声波形辞書に登録する音声波形を選定する登録波形選定手段とを備えることを特徴とする請求項 4 に記載の音声波形辞書作成装置。

【請求項 8】 前記登録波形選定手段は、前記候補波形抽出手段で抽出された全ての前記音声波形のスペクトル情報を抽出するスペクトル抽出手段と、前記スペクトル情報を用いて、前記候補波形抽出手段で抽出された前記音声波形間のスペクトル距離を計算するスペクトル距離計算手段と、前記スペクトル距離に基づいて、前記音声波形辞書に代表として登録する前記音声波形を選定するスペクトル代表波形選定手段とを備えることを特徴とする請求項 7 に記載の音声波形辞書作成装置。

【請求項 9】 前記登録波形選定手段は、前記候補波形抽出手段で抽出された全ての前記音声波形の基本周波数情報を抽出する基本周波数抽出手段と、前記基本周波数情報を用いて、前記候補波形抽出手段で抽出された前記音声波形間の基本周波数距離を計算する基本周波数距離計算手段と、前記基本周波数距離に基づいて、前記音声波形辞書に代表として登録する前記音声波形を選定する基本周波数代表波形選定手段とを備えることを特徴とする請求項 7 に記載の音声波形辞書作成装置。

【請求項 10】 音声合成に使用される音声波形辞書に対する要求条件を入力する辞書条件入力手段と、ネットワークを介して、前記要求条件を送信し、前記音声波形辞書を受信する条件送信辞書受信手段と、前記音声波形辞書を記録する音声波形辞書記録手段と、前記音声波形辞書記録手段に記録された前記音声波形辞書を用いて音声合成を行う音声合成手段と、前記音声合成手段で音声

合成した音声信号を出力する音声出力手段とを備えたことを特徴とする音声合成端末装置。

【請求項11】 前記音声波形辞書記録手段は、複数の前記音声波形辞書を記録することを特徴とする請求項10に記載の音声合成端末装置。

【請求項12】 前記音声合成手段は、複数の音声合成器を備え、複数の前記音声合成器の中から前記音声波形辞書の辞書形式に適した音声合成器を使用して音声合成することを特徴とする請求項10に記載の音声合成端末装置。

【請求項13】 前記音声合成手段は、音声読み上げ対象であるテキスト情報を表音文字列に変換するテキスト解析手段と、前記音声波形辞書から前記表音文字列が一致する前記音声波形を選択する音声波形選択手段と、合成音声の韻律パターンを決定する韻律生成手段と、前記韻律パターンに基づいて前記音声波形の韻律変形を行う音声波形変形手段と、韻律変形された前記音声波形を接続する音声波形接続手段とを備えることを特徴とする請求項10に記載の音声合成端末装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、配信されたテキスト文章を音声読み上げる携帯電話やモバイルコンピュータなどの音声合成端末装置と、音声合成端末装置の利用者の好みに応じた音声波形辞書を作成して音声合成端末装置に送信する音声波形辞書作成装置と、音声合成端末装置及び音声波形辞書作成装置から構成される音声波形辞書配信システムとに関するものである。

【0002】

【従来の技術】従来、電子メールの受信可能な音声合成端末装置としては、特開平11-202885号公報に記載されているように、テキスト情報を音声合成によって読み上げるものが知られている。

【0003】

【発明が解決しようとする課題】しかしながら、上記従来の音声合成端末措置においては、予め搭載されている話者の音声でしか音声合成することができず、利用者の要求する音声で読み上げることができないという問題があった。利用者の要求する話者の音色で合成するためには、例えば予め複数の話者の音声データを音声合成端末措置に所有することが考えられるが、実現するためには膨大な音声波形辞書のデータが必要となり、音声合成端末措置の小型化、低価格化などに大きな障害となる。また、声質変換などの技術を利用することが考えられるが、音色を自由に変換できるインターフェースが提供できたとしても、音声合成端末措置の利用者が直接音響パラメータを修正するのは容易な作業ではなく、加えて現状の声質変換技術では十分に音声合成端末措置の利用者の満足する声質への変換はできない。

【0004】一方、サーバから端末措置へ通信回線を利

用して音声情報を配信する情報サービスでは、電子メールなどのテキスト文章をサーバ側で音声合成し、この合成した音声信号そのものを通信回線により端末措置に配信する方式が考えられる。この場合、サーバ側では複数の話者の音声データを予め用意しておき、端末装置の利用者からの要求のあった話者の音声データでテキスト情報を音声合成して配信すればよく、端末装置側では配信されてきた音声信号をそのまま出力することができる。ところが、このようにサーバ側で音声合成を行って端末装置に送信する方式では、音声読み上げの度に通信回線を接続状態にしておく必要があり、音声信号の伝送に時間がかかることから通信回線の利用時間が長くなり、端末装置の利用者にとっては好ましくない。また、配信時間を短縮するために音声信号を高圧縮し、データ量を減らして配信した場合、圧縮処理により音質劣化が生じて、高音質な音声を提供できなくなってしまう。

【0005】そこで、本発明は、利用者の要求する話者の音色で電子メールなどのテキスト情報を読み上げることが可能な音声波形辞書配信システム、音声波形辞書作成装置、及び音声合成端末装置を提供することを目的とする。

【0006】

【課題を解決するための手段】上記課題を解決するために、本発明の音声波形辞書配信システムは、音声合成に使用される音声波形辞書を作成する音声波形辞書作成装置と、前記音声波形辞書をネットワーク経由で取得して音声合成を行う音声合成端末装置とから構成され、前記音声波形辞書作成装置は、複数の音声データベースと、前記音声合成端末装置から受信する前記音声波形辞書に対する要求条件に基づいて、前記音声波形辞書の作成条件を決定する辞書条件決定手段と、前記音声データベースの中から前記作成条件に基づいた音声データベースを用いて前記音声波形辞書を作成する音声波形辞書作成手段と、前記ネットワークを介して、前記要求条件を受信し、前記音声波形辞書を送信する条件受信辞書送信手段とを備え、前記音声合成端末装置は、前記要求条件を入力する辞書条件入力手段と、前記ネットワークを介して、前記要求条件を送信し、前記音声波形辞書を受信する条件送信辞書受信手段と、前記音声波形辞書を記録する音声波形辞書記録手段と、前記音声波形辞書記録手段に記録された前記音声波形辞書を用いて音声合成を行う音声合成手段と、前記音声合成手段で音声合成した音声信号を出力する音声出力手段とを備える構成を有している。

【0007】この構成により、本発明の音声波形辞書配信システムは、音声波形辞書作成装置に複数種類の話者の音声データベースを用意しておき、音声合成端末装置の利用者が要求する音声波形辞書のサイズや音色の音声波形辞書を音声波形辞書作成装置で自由に作成し、音声波形辞書作成装置から音声合成端末装置にダウンロード

できるので、音声合成端末装置の音声合成器で好みの話者の音色で電子メールなどのテキスト情報を音声読み上げすることが可能となる。

【0008】また、本発明の音声波形辞書配信システムは、音声波形辞書作成装置の管理者が音声データベースの追加や修正を行い、音声合成端末装置の利用者が最新の音声波形辞書をダウンロードすることによって、音声合成端末装置の音声波形辞書を容易にメンテナンスすることができる。

【0009】また、本発明の音声波形辞書配信システムは、前記音声波形辞書作成装置は、前記音声波形辞書作成手段で作成した前記音声波形辞書を用いて音声合成を実行する作成辞書音声合成手段を備え、前記音声合成端末装置は、前記ネットワークを介して、前記作成辞書音声合成手段で音声合成した音声信号を受信し、受信した前記音声信号を前記音声出力手段で出力する構成を有している。

【0010】この構成により、本発明の音声波形辞書配信システムは、音声波形辞書作成装置で作成した音声波形辞書を音声合成端末装置にダウンロードする前に、音声合成端末装置の利用者に、音声波形辞書作成装置で作成した音声波形辞書を使って音声合成した試聴用の音声信号を試聴させることができ、作成した音声波形辞書をダウンロードするか否かを選択させることが可能となる。

【0011】また、本発明の音声波形辞書作成装置は、複数の音声データベースと、音声合成に使用される音声波形辞書に対する要求条件に基づいて、前記音声波形辞書の作成条件を決定する辞書条件決定手段と、前記音声データベースの中から前記作成条件に基づいた音声データベースを用いて前記音声波形辞書を作成する音声波形辞書作成手段と、ネットワークを介して、前記要求条件を受信し、前記音声波形辞書を送信する条件受信辞書送信手段とを備えた構成を有している。

【0012】この構成により、本発明の音声波形辞書作成装置は、予め複数種類の話者の音声データベースを用意しておき、音声合成端末装置の利用者が要求する音声波形辞書のサイズや音色の音声波形辞書を自由に作成し、音声合成端末装置に配信することができるので、音声合成端末装置の音声合成器に、音声合成端末装置の利用者が要求する話者の音色で電子メールなどのテキスト情報を音声読み上げさせることが可能となる。

【0013】また、本発明の音声波形辞書作成装置は、前記音声波形辞書作成手段は、前記音声データベースに予め付属させられた付属情報を参照して音声波形を選出する音声波形選出手段と、前記音声波形選出手段によって選出された前記音声波形を所定の形式に変換するパラメータ化手段と、前記音声データベースから前記音声波形に付属させられた前記付属情報を抽出する音声波形付属情報抽出手段と、前記パラメータ化手段でパラメータ

化された前記音声波形、及び、前記音声波形付属情報抽出手段によって抽出された前記付属情報を用いて、前記音声波形辞書を構築する辞書構築手段とを備える構成を有している。

【0014】この構成により、本発明の音声波形辞書作成装置は、例えば音声合成端末装置で用いられている合成方式がLPC系のパラメトリックな合成方式である場合や波形重畳型の合成方式などである場合、それぞれの合成方式に応じて指定された形式の音声波形辞書を作成することができ、音声合成端末装置に搭載されている音声合成器の種類に関わらず、音声合成端末装置の利用者が要求する話者の音声波形辞書を作成し、音声合成端末装置に配信することが可能である。

【0015】また、本発明の音声波形辞書作成装置は、前記パラメータ化手段は、前記音声波形を音源特性と声道特性とに分離して、それぞれを独立にパラメータ化する構成を有している。

【0016】この構成により、本発明の音声波形辞書作成装置は、音声波形を音源特性と声道特性とに分離することで、音声合成の際には音声の音色などスペクトルに依存する特徴と、声の高さなど音源に依存する特徴とを独立に制御することが容易になる。更に、本発明の音声波形辞書作成装置は、パラメータ化によって音声波形のデータ容量を大幅に圧縮できる音声波形辞書を作成することが可能である。

【0017】また、本発明の音声波形辞書作成装置は、前記パラメータ化手段は、前記音声波形をピッチ波形単位に分割して抽出するピッチ波形抽出手段と、前記ピッチ波形抽出手段によって抽出された前記ピッチ波形に対して、類似度が所定値以上の前記ピッチ波形同士を分類するピッチ波形分類手段と、前記ピッチ波形分類手段によって分類された前記ピッチ波形の中から前記音声波形辞書に代表として登録する代表ピッチ波形を作成する代表ピッチ波形作成手段とを備える構成を有している。

【0018】この構成により、本発明の音声波形辞書作成装置は、例えば合成単位内または合成単位の枠を越えて類似度の高いピッチ波形が存在する場合、これら複数のピッチ波形を代表的な一つのピッチ波形に置き換えることで、無駄の少ない効率的な音声波形辞書を作成することが可能であり、音声波形辞書のデータ容量を削減することができる。

【0019】また、本発明の音声波形辞書作成装置は、前記音声波形選出手段は、予め音声データに付属させられた音素境界ラベル情報を参照して合成単位の境界ラベルを付与し、前記合成単位を作成する合成単位作成手段と、前記合成単位作成手段で作成された前記合成単位の前記音声波形を前記音声波形辞書に登録する候補として抽出する候補波形抽出手段と、前記候補波形抽出手段で抽出された前記音声波形の中から前記音声波形辞書に登録する音声波形を選定する登録波形選定手段とを備える

構成を有している。

【0020】この構成により、本発明の音声波形辞書作成装置は、音声合成端末装置の利用者に合成単位を自由に指定させることができ、例えば短い合成単位を指定されればサイズの小さな音声波形辞書を作成することができ、長い合成単位を指定されれば、高音質な音声合成が可能な音声波形辞書を作成することができる。したがって、本発明の音声波形辞書作成装置によれば、音声合成端末装置の利用者は、音声波形辞書を記録するメモリ容量と、合成音声の音質との兼ね合いを判断して音声波形辞書を作成することが可能となる。

【0021】また、本発明の音声波形辞書作成装置は、前記登録波形選定手段は、前記候補波形抽出手段で抽出された全ての前記音声波形のスペクトル情報を抽出するスペクトル抽出手段と、前記スペクトル情報を用いて、前記候補波形抽出手段で抽出された前記音声波形間のスペクトル距離を計算するスペクトル距離計算手段と、前記スペクトル距離に基づいて、前記音声波形辞書に代表として登録する前記音声波形を選定するスペクトル代表波形選定手段とを備える構成を有している。

【0022】この構成により、本発明の音声波形辞書作成装置は、音声合成端末装置の利用者の要求する音声波形辞書のサイズに応じた音声波形数で、音声合成時に接続する二つの合成単位間で生じるスペクトル歪を少なくすることが可能な音声波形辞書を作成することができる。

【0023】また、本発明の音声波形辞書作成装置は、前記登録波形選定手段は、前記候補波形抽出手段で抽出された全ての前記音声波形の基本周波数情報を抽出する基本周波数抽出手段と、前記基本周波数情報を用いて、前記候補波形抽出手段で抽出された前記音声波形間の基本周波数距離を計算する基本周波数距離計算手段と、前記基本周波数距離に基づいて、前記音声波形辞書に代表として登録する前記音声波形を選定する基本周波数代表波形選定手段とを備える構成を有している。

【0024】この構成により、本発明の音声波形辞書作成装置は、音声合成端末装置の利用者の要求する音声波形辞書サイズに応じた音声波形数で、ピッチ変換や時間長変換などの韻律変換処理によって生じる波形歪を抑えることが可能な音声波形辞書を作成することができる。

【0025】また、本発明の音声合成端末装置は、音声合成に使用される音声波形辞書に対する要求条件を入力する辞書条件入力手段と、ネットワークを介して、前記要求条件を送信し、前記音声波形辞書を受信する条件送信辞書受信手段と、前記音声波形辞書を記録する音声波形辞書記録手段と、前記音声波形辞書記録手段に記録された前記音声波形辞書を用いて音声合成を行う音声合成手段と、前記音声合成手段で音声合成した音声信号を出力する音声出力手段とを備えた構成を有している。

【0026】この構成により、本発明の音声合成端末装

置は、利用者が音声波形辞書を決定することができ、情報提供サーバから獲得したテキスト情報を利用者の好みの話者の声で音声合成を行うことができる。また、本発明の音声合成端末装置を使用した音声波形辞書配信システムは、サーバ（音声波形辞書作成装置）側で音声合成して音声信号を送受信するシステムと比較して、音声波形辞書を一度ダウンロードすれば音声合成端末装置に音声信号を通信する必要がないため、通信負荷を抑えることができる。

10 【0027】また、本発明の音声合成端末装置は、前記音声波形辞書記録手段は、複数の前記音声波形辞書を記録する構成を有している。

【0028】この構成により、本発明の音声合成端末装置は、一つのアプリケーションの中で複数話者での合成が可能となり、例えば電子メールやニュース文などコンテンツ毎に異なる話者での音声読み上げが可能となるので、利用者は現在の音声読み上げの対象が何であるのかを判断することができる。

20 【0029】また、本発明の音声合成端末装置は、前記音声合成手段は、複数の音声合成器を備え、複数の前記音声合成器の中から前記音声波形辞書の辞書形式に適した音声合成器を使用して音声合成する構成を有している。

30 【0030】この構成により、本発明の音声合成端末装置は、例えば音声波形辞書の辞書形式がLPCパラメータである場合や、音声波形をピッチ波形単位に分割した形式である場合など、それぞれの辞書形式に応じて音声合成器を自動的に使い分けることで、音声波形辞書をダウンロードする際、利用者に辞書形式を把握させる必要がない。

【0031】また、本発明の音声合成端末装置は、前記音声合成手段は、音声読み上げ対象であるテキスト情報を表音文字列に変換するテキスト解析手段と、前記音声波形辞書から前記表音文字列が一致する前記音声波形を選択する音声波形選択手段と、合成音声の韻律パターンを決定する韻律生成手段と、前記韻律パターンに基づいて前記音声波形の韻律変形を行う音声波形変形手段と、韻律変形された前記音声波形を接続する音声波形接続手段とを備える構成を有している。

40 【0032】この構成により、本発明の音声合成端末装置は、音声波形辞書に登録されている複数の音声波形の候補の中から、合成環境に最も適した音声波形を選択することができるようになり、高音質な音声合成が可能となる。

【0033】

【発明の実施の形態】以下、本発明の実施の形態について、図面を用いて説明する。

（第1の実施の形態）

50 【0034】まず、第1の実施の形態に係る音声波形辞書配信システムの構成について説明する。

【0035】図1において、本実施の形態に係る音声波形辞書配信システム100は、音声合成に用いられる音声波形辞書を作成する音声波形辞書作成装置110と、有線、無線を問わず、データの通信が可能な通信回線によって構成されるネットワーク120と、音声波形辞書作成装置110によって作成された音声波形辞書をネットワーク120を介して取得し、取得した音声波形辞書を用いて音声合成を行う音声合成端末装置130、音声合成端末装置140及び音声合成端末装置150と、電子メールやニュース文などのコンテンツをテキスト情報

としてネットワーク120を介して音声合成端末装置130、音声合成端末装置140及び音声合成端末装置150に提供する情報提供サーバ160とを備えている。

【0036】また、音声波形辞書作成装置110は、音声合成端末装置130の利用者が要求する話者の種類や音声波形辞書の形式、サイズなどの情報を受け取って音声波形辞書の作成条件を決定する辞書条件決定部（辞書条件決定手段）111と、音声データベース112a、音声データベース112b及び音声データベース112cなどの複数の音声データベースから構成される音声データベース群112と、辞書条件決定部111によって決定された作成条件に基づいて、音声データベース群112内の音声データベースを加工して音声波形辞書を作成する音声波形辞書作成部（音声波形辞書作成手段）113と、ネットワーク120を介して音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150とのデータ通信を制御し、音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150の利用者が要求する話者の種類や音声波形辞書の形式、サイズなどの情報を受信して辞書条件決定部111に渡したり、音声波形辞書作成部113によって作成された音声波形辞書を音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150に送信したりするデータ送受信部（条件受信辞書送信手段）114とを備えている。

【0037】なお、音声データベース112a、音声データベース112b及び音声データベース112cなどの音声データベースは、音声波形と、音素境界のラベリングやピッチマークなどの付属情報とによって話者毎に構成されている。例えば、音声データベース112aは、話者1に、音声データベース112bは、話者2に、音声データベース112cは、話者3に対応している。ここで、音声データベース112a、音声データベース112b及び音声データベース112cなどの音声データベースは、音声波形として試聴用の音声波形も備えている。

【0038】また、音声合成端末装置130は、話者の種類や音声波形辞書のサイズなどの音声波形辞書の要求条件を利用者によって入力されるキーボードやマウスなどの辞書条件入力部（辞書条件入力手段）131と、受

け取った1つ又は複数の音声波形辞書を記録する書き換えが自由な記録媒体（音声波形辞書記録手段）132と、記録媒体132によって記録された音声波形辞書を用いて音声合成を行う複数の音声合成器によって構成される音声合成部（音声合成手段）133と、音声合成部133によって音声合成された音声信号や、後述するデータ送受信部135で受信した音声信号を出力するスピーカ（音声出力手段）134と、ネットワーク120を介して音声波形辞書作成装置110や情報提供サーバ160などの外部装置とのデータ通信を制御し、利用者が要求する話者の種類や音声波形辞書の形式、サイズなどの情報を音声波形辞書作成装置110に送信したり、音声波形辞書作成装置110や情報提供サーバ160から音声波形辞書やテキスト情報などを受信したりするデータ送受信部（条件送信辞書受信手段）135とを備えている。

【0039】なお、音声合成端末装置140及び音声合成端末装置150は、音声合成端末装置130と同様な構成であるので詳細な説明を省略する。

【0040】次に、本実施の形態に係る音声波形辞書配信システムの動作について説明する。

【0041】音声合成端末装置130の利用者は、図示していない入力部への入力によって、音声波形辞書作成装置110の音声データベース群112の音声データベースの中からサンプル音声を試聴する音声データベースを選択する。ここで、音声合成端末装置130の利用者から入力部への入力は、信号として、データ送受信部135及びネットワーク120を介して、音声波形辞書作成装置110のデータ送受信部114で受信される。

【0042】そして、データ送受信部114は、受信した信号に応じて、音声データベース群112の中の利用者によって選択された音声データベースから試聴用の音声波形（音声信号）を取得し、ネットワーク120を介して、音声合成端末装置130のデータ送受信部135に送信する。データ送受信部135に送信された音声信号は、データ送受信部135によってスピーカ134に出力され、スピーカ134によってサンプル音声として出力される。

【0043】以上のようにして、音声合成端末装置130の利用者は、音声波形辞書作成装置110の音声データベース群112の音声データベースの中からサンプル音声を試聴し、自分の気に入った話者を探すことができる。

【0044】そして、音声合成端末装置130の利用者は、音声波形辞書作成装置110の音声データベース群112の音声データベースの中に音声読み上げに使用することを要求する話者が存在した場合、話者の種類や音声波形辞書のサイズなどの音声波形辞書に対する要求条件を辞書条件入力部131から入力することができる。

【0045】音声合成端末装置130の利用者が、話者

の種類や音声波形辞書のサイズなどの音声波形辞書の要求条件を辞書条件入力部131から入力すると、音声合成端末装置130は、辞書条件入力部131から入力された音声波形辞書の要求条件をデータ送受信部135からネットワーク120を介して音声波形辞書作成装置110に送信する。

【0046】また、音声合成端末装置130は、音声合成部133で使用している音声合成器の種類や、記録媒体132で使用可能なメモリ残量情報などの音声波形辞書の使用可能条件も、データ送受信部135からネットワーク120を介して自動的に音声波形辞書作成装置110に送信する。

【0047】音声波形辞書作成装置110は、音声合成端末装置130から送信された音声波形辞書の要求条件及び使用可能条件をデータ送受信部114によって受信すると、受信した音声波形辞書の要求条件及び使用可能条件に基づいて、辞書条件決定部111によって音声合成端末装置130の利用者がダウンロードできる音声波形辞書の作成条件を決定し、決定した音声波形辞書の作成条件を辞書条件決定部111から音声波形辞書作成部113に出力する。

【0048】音声波形辞書作成部113は、辞書条件決定部111から出力される音声波形辞書の作成条件に基づいて、音声データベース群112内の音声データベースを加工して音声波形辞書を作成し、作成した音声波形辞書をデータ送受信部114に出力する。

【0049】より詳細に説明すると、音声波形辞書作成部113では、例えば同じ話者の音声波形辞書でも、音声合成に必要な最低限の音韻並びの音声波形しか持たない標準的な音声波形辞書や、調音結合の影響や韻律パターンの種類も考慮した音声波形を用いる最高音質版の合成が可能な音声波形辞書など、音声合成端末装置130の利用者の要求する音質や音声波形辞書のサイズに応じて音声波形辞書をカスタマイズすることができる。

【0050】また、音声波形辞書作成部113では、音声合成端末装置130で使用されている音声合成器に合った音声波形辞書を作成することができる。例えば、音声合成端末装置130の音声合成部133で使用されている音声合成器が波形重畳型合成（PSOLA）方式である場合は、ピッチ波形配列の羅列によって音声波形辞書が構成され、LSPなどに代表されるLPC系のパラメトリックな合成方式である場合は、フレーム毎のパラメータ系列によって音声波形辞書が構成される。

【0051】なお、音声波形辞書作成部113は、上述したように音声合成端末装置130の利用者の要求に応じてその場で音声波形辞書を作成し、作成した音声波形辞書をデータ送受信部114に出力するようにしても良いし、予め複数の話者の音声波形辞書を作成して用意しておいて、用意しておいた音声波形辞書のうち、音声合成端末装置130の利用者の要求に応じた音声波形辞書

をデータ送受信部114に出力するようにしても良い。

【0052】音声波形辞書作成装置110は、音声波形辞書作成部113からデータ送受信部114に音声波形辞書を出力されると、音声波形辞書作成部113から出力された音声波形辞書をデータ送受信部114からネットワーク120を介して音声合成端末装置130に送信する。

【0053】音声合成端末装置130は、音声波形辞書作成装置110から送信された音声波形辞書をデータ送受信部135によって受信すると、受信した音声波形辞書を記録媒体132によって格納する。

【0054】以上のようにして、音声合成端末装置130の利用者は、音声波形辞書作成装置110の音声データベース群112の音声データベースの中に音声読み上げに使用することを要求する話者が存在した場合、その話者の音声波形辞書を音声波形辞書作成装置110で作成して自分の音声合成端末装置130にダウンロードすることができる。

【0055】したがって、音声合成端末装置130では、情報提供サーバ160などからネットワーク120を介してデータ送受信部135で受信したテキスト情報を音声合成部133に出力し、音声合成部133において記録媒体132に格納された音声波形辞書を使用して音声合成することによって、音声読み上げを可能とする。

【0056】また、音声合成端末装置130は、一度ダウンロードした音声波形辞書をずっと使い続けることも可能だし、何度でも書き換えることも可能である。また、音声合成端末装置130は、記録媒体132のメモリ容量に応じて、複数の音声波形辞書を保持することも可能である。

【0057】なお、音声合成端末装置140及び音声合成端末装置150の動作については、音声合成端末装置130の動作と同様であるので詳細な説明を省略する。

【0058】以上説明したように、音声波形辞書配信システム100は、音声波形辞書作成装置110に複数種類の話者の音声データベースを用意しておき、音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150の利用者が要求する音声波形辞書のサイズや音色の音声波形辞書を音声波形辞書作成装置110で自由に作成し、音声波形辞書作成装置110から音声合成端末装置130にダウンロードできるので、音声合成端末装置130の音声合成器で好みの話者の音色で電子メールなどのテキスト情報を音声読み上げすることが可能となる。

【0059】また、音声波形辞書配信システム100は、音声波形辞書作成装置110の管理者が音声データベース群112に対して音声データベースの追加や修正を行い、音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150の利用者が最新の音声

波形辞書をダウンロードすることによって、音声合成端末装置130、音声合成端末装置140又は音声合成端末装置150の音声波形辞書を容易にメンテナンスすることができる。

(第2の実施の形態)

【0060】まず、第2の実施の形態に係る音声波形辞書配信システムの構成について説明する。なお、本実施の形態に係る音声波形辞書配信システムの構成のうち、第1の実施の形態に係る音声波形辞書配信システムの構成と同様な構成については、同一の符号を付して詳細な説明を省略する。

【0061】図2において、本実施の形態に係る音声波形辞書配信システム200は、第1の実施の形態に係る音声波形辞書配信システムの音声波形辞書作成装置110(図1参照)の代わりに、音声波形辞書作成装置201を備えている。

【0062】音声波形辞書作成装置201は、辞書条件決定部111と、音声データベース112a、音声データベース112b及び音声データベース112cなどの複数の音声データベースから構成される音声データベース群112と、音声波形辞書作成部113と、データ送受信部114と、音声波形辞書作成部113によって作成された音声波形辞書を用いて音声合成を行う音声合成器から構成され、試聴用の音声信号を合成する音声合成部(作成辞書音声合成手段)202とを備えている。

【0063】次に、本実施の形態に係る音声波形辞書配信システムの動作について説明する。

【0064】図3において、第1の実施の形態と同様にして、音声合成端末装置130の利用者が、音声合成端末装置130の辞書条件入力部131から音声波形辞書の要求条件を設定すると(ステップS301)、音声波形辞書作成装置201は、音声合成端末装置130の利用者によって設定された要求条件や、音声合成端末装置130から受信する使用可能条件に基づいて音声波形辞書の作成条件を決定し、決定した作成条件に基づいて音声波形辞書作成部113によって音声波形辞書を作成する(ステップS302)。

【0065】音声波形辞書作成装置201は、音声波形辞書作成部113によって音声波形辞書を作成すると、音声合成部202によって、作成した音声波形辞書を用いて試聴用の音声信号の音声合成を行う。そして、音声合成された音声信号は、音声合成部202から、データ送受信部114、ネットワーク120、及び、音声合成端末装置130のデータ送受信部135を介して、音声合成端末装置130のスピーカ134から出力される。

【0066】したがって、音声合成端末装置130の利用者は、スピーカ134から出力される試聴用の音声信号を試聴することができ(ステップS303)、試聴した音声信号に基づいて、音声波形辞書作成装置201によって作成した音声波形辞書をダウンロードするか否か

を決定することができる(ステップS304)。

【0067】音声合成端末装置130の利用者は、音声波形辞書作成装置201によって作成した音声波形辞書をダウンロードすると決定すると、ダウンロード命令を図示していない入力部に入力する。入力されたダウンロード命令は、音声合成端末装置130からネットワーク120を介して音声波形辞書作成装置201に送信される。音声波形辞書作成装置201に音声合成端末装置130からダウンロード命令が送信されると、音声波形辞書作成部113によって作成された音声波形辞書は、第1の実施の形態と同様にして、音声波形辞書作成部113から、データ送受信部114、ネットワーク120、及び、音声合成端末装置130のデータ送受信部135を介して、音声合成端末装置130の記録媒体132に送信されて、格納される(ステップS305)。

【0068】一方、音声合成端末装置130の利用者は、音声波形辞書作成装置201によって作成した音声波形辞書をダウンロードしないと決定する場合、他の要求条件で音声辞書を作成し直すか否かを判断し(ステップS306)、他の要求条件で音声波形辞書を作成し直すと判断したときには、ステップS301の処理から繰り返す。

【0069】以上説明したように、音声波形辞書配信システム200は、音声波形辞書作成装置201で作成した音声波形辞書を音声合成端末装置130にダウンロードする前に、音声合成端末装置130の利用者に、音声波形辞書作成装置201で作成した音声波形辞書を使って音声合成した試聴用の音声信号を試聴させることができるので、作成した音声波形辞書をダウンロードするか否かを選択させることが可能となる。

【0070】なお、第1の実施の形態に係る音声波形辞書配信システムで出力されるサンプル音声は、音声データベース群112の音声データベースの試聴用の音声波形をそのまま出力する原音声であったのに対し、音声波形辞書配信システム200で出力されるサンプル音声は、音声波形辞書作成装置201で作成した音声波形辞書を使って音声合成した合成音声である。したがって、第1の実施の形態に係る音声波形辞書配信システムが、試聴によって音声データベース間の話者の音色の違いを比較することができるのに対し、音声波形辞書配信システム200は、試聴によって、音声データベース間の話者の音色の違いに加え、音声波形辞書のデータ圧縮方法や合成方式に起因する音質の違いなども比較することができる。

(第3の実施の形態)

【0071】図4に示すように、第3の実施の形態に係る音声波形辞書作成装置401の構成は、第1の実施の形態に係る音声波形辞書配信システムの音声波形辞書作成装置の構成と同様であるので、同一の符号を付して詳細な説明を省略する。

【0072】また、本実施の形態に係る音声波形辞書作成装置401は、第1の実施の形態に係る音声波形辞書配信システムの音声波形辞書作成装置と同様に動作することができる。

【0073】したがって、音声波形辞書作成装置401は、予め複数種類の話者の音声データベースを用意しておき、音声合成端末装置の利用者が要求する音声波形辞書のサイズや音色の音声波形辞書を自由に作成し、音声合成端末装置に配信することができるので、音声合成端末装置の音声合成器に、音声合成端末装置の利用者が要求する話者の音色で電子メールなどのテキスト情報を音声読み上げさせることが可能となる。

(第4の実施の形態)

【0074】まず、第4の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第3の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であるので、第3の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符号を付して詳細な説明を省略する。

【0075】図5に示すように、本実施の形態に係る音声波形辞書作成装置は、第3の実施の形態に係る音声波形辞書作成装置の音声波形辞書作成部113(図4参照)の代わりに、音声波形辞書作成部(音声波形辞書作成手段)501を備えている。

【0076】音声波形辞書作成部501は、音声データベース112aの中から、予めラベリングされたラベル情報を参照して辞書に含める音声波形を選出する音声波形選出部(音声波形選出手段)502と、音声波形選出部502で選出された音声波形を、辞書条件決定部111によって決定された所定の形式に変換するパラメータ化部(パラメータ化手段)503と、音声波形選出部502で選出された音声波形に対応する付属情報を音声データベース112aから抽出し、必要な付属情報を音声波形辞書に格納する音声波形付属情報抽出部(音声波形付属情報抽出手段)504と、パラメータ化部503で所定の形式に変換された音声波形、及び、音声波形付属情報抽出部504で抽出された付属情報を書き出し、音声波形辞書を構築する辞書構築部(辞書構築手段)505とを備えている。

【0077】なお、音声データベース112aなどの音声データベースは、上述したように話者毎に音声波形を有しており、更に、音声波形の抽出環境を示すコンテキスト情報や音素・音節境界ラベル、基本周波数F0、ピッチマークなどの音声波形に対応する付属情報を有している。

【0078】次に、本実施の形態に係る音声波形辞書作成装置の動作について説明する。

【0079】まず、辞書条件決定部111では、音声合成端末装置(図示していない)の利用者の要求する辞書

サイズに応じて、合成単位を決定する。合成単位としては、例えば音素、音節、VCV(母音—子音—母音)連鎖などが考えられる。また、音声合成を行うのに必要最低限の音声波形だけを音声波形辞書に格納するのか、それとも同じ音韻並びの音声波形を複数用意するのかなども、辞書条件決定部111で決定される。

【0080】次いで、音声波形選出部502では、辞書条件決定部111で決定された合成単位や音声波形数などに基づき、音声データベース112aに含まれているラベル情報を参照して必要な音声波形を選出する。もし同じ音韻並びの音声波形に複数の候補がある場合は、音声波形の韻律パターンや前後の音韻環境などを考慮して音声波形を選出する。

【0081】次いで、パラメータ化部503では、音声波形選出部502によって選出された音声波形を、音声合成端末装置で使用されている音声合成器に合わせて加工する。例えば音声合成端末装置で使用されている音声合成器がPSOLA法に基づく方式である場合は、音声波形をピッチ波形単位に分割し、このピッチ波形をパラメータとする。また、例えば音声合成器がLPC系の方式である場合は、音声波形をフレーム分割して線形予測分析を行い、パラメータ化する。

【0082】次いで、辞書構築部505では、パラメータ化部503で所定の形式に変換された音声波形を書き出し、音声波形辞書を構築する。なお、辞書構築部505は、音声波形辞書を構築する際、音声波形付属情報抽出部504で抽出された各音声波形の付属情報を辞書に付属させる。ここで、付属情報は、上述したように音素並びや基本周波数パターン、音素境界ラベルなどであり、音声合成の際に音声波形選択のために参照される。

【0083】最後に、辞書構築部505で構築された音声波形辞書は、第1の実施の形態において説明したように、データ送受信部114を経由して音声合成端末装置に送信される。

【0084】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、例えば音声合成端末装置で用いられている合成方式がLPC系のパラメトリックな合成方式である場合や波形重畳型の合成方式などである場合、それぞれの合成方式に応じて指定された形式の音声波形辞書を作成することができ、音声合成端末装置に搭載されている音声合成器の種類に関わらず、音声合成端末装置の利用者が要求する話者の音声波形辞書を作成し、音声合成端末装置に配信することが可能である。

(第5の実施の形態)

【0085】まず、第5の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であるので、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符号を付し

て詳細な説明を省略する。

【0086】図6に示すように、本実施の形態に係る音声波形辞書作成装置は、第4の実施の形態に係る音声波形辞書作成装置のパラメータ化部503（図5参照）の代わりに、パラメータ化部（パラメータ化手段）601を備えている。

【0087】パラメータ化部601は、音声波形を、主に音声の高さや強さなどを決定するパラメータである音源情報（音源特性）と、主に音声の音韻の種類や音色などに寄与するパラメータであるスペクトル情報（声道特性）とに分離する音声分析を行う音源・声道特性分離処理部602を備えている。

【0088】次に、本実施の形態に係る音声波形辞書作成装置の動作について説明する。

【0089】音声波形選出部502で選出された音声波形は、パラメータ化部601によって、音声合成器で採用されている音声合成方式に適した形式にパラメータ化が行なわれる。

【0090】ここで、パラメータ化部601では、音源・声道特性分離処理部602によって、音声波形611を音源情報とスペクトル情報とに分離する。例として音源・声道特性分離処理部602に準同型分析が用いられる場合、スペクトル情報はスペクトル包絡612、音源情報は音源パルス列613に分離される。

【0091】このように音声波形を音源・声道特性分離処理部602でスペクトル情報と音源情報とに分離することで、主に音韻の種類や音色に特徴を与えるスペクトル情報と、声の高さや強さを制御する音源情報とを、それぞれ独立に制御することができるようになる。ここで分離されたスペクトル情報と音源情報とは、それぞれの重要度に応じて情報圧縮が行なわれ、辞書構築部505で音声波形辞書に格納される。

【0092】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、音声波形を音源特性と声道特性とに分離することで、音声合成の際には音声の音色などスペクトルに依存する特徴と、声の高さなど音源に依存する特徴とを独立に制御することが容易になる。更に、本実施の形態に係る音声波形辞書作成装置は、パラメータ化によって音声波形のデータ容量を大幅に圧縮できる音声波形辞書を作成することが可能である。

【0093】なお、本実施の形態においては、準同型分析によって音源特性と声道特性を分離する例を示したが、本発明によれば、線形予測分析が用いられる場合は、声道特性として線形予測係数、音源特性として残差信号に分離することも可能である。

（第6の実施の形態）

【0094】まず、第6の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であ

るので、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符合を付して詳細な説明を省略する。

【0095】図7に示すように、本実施の形態に係る音声波形辞書作成装置は、第4の実施の形態に係る音声波形辞書作成装置のパラメータ化部503（図5参照）の代わりに、パラメータ化部（パラメータ化手段）701を備えている。

【0096】パラメータ化部701は、音声波形選出部502で音声波形辞書に登録することが決定された音声波形について、ピッチマークを基準に抽出窓を用いてピッチ波形単位に分解するピッチ波形抽出部（ピッチ波形抽出手段）702と、ピッチ波形抽出部702で分解されたピッチ波形を類似度を評価尺度としてクラスタリングし、類似度が所定値以上のピッチ波形（即ち、波形形状の似ているピッチ波形）同士をまとめあげるピッチ波形分類部（ピッチ波形分類手段）703と、ピッチ波形分類部703でのクラスタリングによって生成された各クラスタの中から、音声波形辞書に代表として登録する代表ピッチ波形を選出する代表ピッチ波形作成部（代表ピッチ波形作成手段）704とを備えている。

【0097】また、音声波形選出部502は、音声データベースの中から音声波形辞書に登録する音声波形を韻律バランスやスペクトル距離などを考慮して選出するようになっており、辞書構築部505は、音声波形選出部502で選出された音声波形によって音声波形辞書を構築するようになっている。

【0098】ただし、辞書構築部505において実際に音声波形辞書に書き出されるのは、後述するように、各音声波形を構成し、パラメータ化部701によって選出される代表ピッチ波形である。

【0099】次に、本実施の形態に係る音声波形辞書作成装置の動作について説明する。

【0100】図8に示すように、音声波形選出部502で選出されたVCV単位の音声波形801～804は、音声波形辞書に登録されることになる。

【0101】ここで、音声合成端末装置（図示していない）で採用されている音声合成器が波形重畳型の音声合成方式を採用している場合、音声波形801～804のうち有声音はピッチ波形抽出部702で1ピッチ周期相当分の波形（ピッチ波形）に区切られる。このピッチ波形は予め音声波形に付属させられたピッチマークを基準に、ハニング窓を用いて抽出される。

【0102】ピッチ波形は周期性のある音声波形から抽出された波形であるため、ピッチ波形間で類似度の高いものが多い。そのためピッチ波形分類部703では、ピッチ波形間の類似度を評価関数として波形形状が似ていて類似度が所定値以上のピッチ波形同士をまとめてクラスタ811及びクラスタ812を作成する。ここで、類似度を表す評価関数は、ピッチ波形間の相関係数やユー

クリッド距離などが用いられる。

【0103】代表ピッチ波形作成部704では、各クラス（クラスタ811及びクラスタ812）のセントロイドに最も近いピッチ波形を代表ピッチ波形821及び代表ピッチ波形822とする。すなわち、同一クラス内の複数のピッチ波形が一つのピッチ波形に置き換えられることになる。音声波形801～804の有声音部分は、もとのピッチ波形の代わりにこの代表ピッチ波形によって構成される。

【0104】辞書構築部505では、音声波形辞書に、代表ピッチ波形と、この代表ピッチ波形を構成要素とする音声波形との関係を登録する。

【0105】したがって、音声合成端末装置130は、音声合成の際に、必要な音声波形と結びついたピッチ波形を用いることができる。

【0106】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、例えば合成単位内または合成単位の枠を越えて類似度の高いピッチ波形が存在する場合、これら複数のピッチ波形を代表的な一つのピッチ波形に置き換えることで、無駄の少ない効率的な音声波形辞書を作成することが可能であり、音声波形辞書のデータ容量を削減することができる。

（第7の実施の形態）

【0107】まず、第7の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であるので、第4の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符号を付して詳細な説明を省略する。

【0108】図9に示すように、本実施の形態に係る音声波形辞書作成装置は、第4の実施の形態に係る音声波形辞書作成装置の音声波形選出部502（図5参照）の代わりに、音声波形選出部（音声波形選出手段）901を備えている。

【0109】音声波形選出部901は、音声データベース112aの中に含まれる音素境界ラベル情報を参照して、指定された合成単位で音声波形を抽出するための区切り情報（合成単位の境界ラベル）を作成する合成単位作成部（合成単位作成手段）902と、合成単位作成部902で作成された区切り情報に基づき音声波形を音声波形辞書に登録する候補として切り出し、切り出した音声波形の一覧をリストにする候補波形抽出部（候補波形抽出手段）903と、候補波形抽出部903で抽出された音声波形の中から、音声波形辞書に登録する音声波形を選定する登録波形選定部（登録波形選定手段）904とを備えている。

【0110】また、辞書条件決定部111は、話者の種類や音声波形辞書の形式に加えて、音声合成端末装置の利用者の所望する合成単位を受け取って音声波形辞書の

作成条件を決定するようになっており、パラメータ化部503は、登録波形選定部904で選定された音声波形を、音声波形辞書作成装置の音声合成器の条件に合致した所定の波形（パラメータ）形式に変換するようになっている。

【0111】また、音声データベース112aは、図10に示すように、付属情報の一つとして、隣接する音素との境界を示し、各音素の終点位置情報とその音素の種類を表す記号である音素境界ラベル1010を格納している。ここで、音素境界ラベル1010は、音声波形1000に予め付属させられており、同様に、音声データベース112aに格納される全ての音声波形には、音素境界ラベルが予め付属させられている。なお、音声波形1000は、自然発生の音声データであり、音声データベース112aを構成する文の一例である。

【0112】次に、本実施の形態に係る音声波形辞書作成装置の動作について説明する。

【0113】辞書条件決定部111は、音声合成端末装置の利用者の要求や、音声合成端末装置の音声合成器の条件によって、合成単位を決定する。ここで、合成単位とは、音声合成時に接続する音声波形の最小単位のことであり、音素単位1011、CV（子音・母音）・VC（母音・子音）単位1012、VCV（母音・子音・母音）単位1013などが一般的に用いられる。

【0114】例えば、音声合成端末装置の音声合成器の条件によって、音素が合成単位として選択された場合、合成単位作成部902は、音声データベース112a内に格納されている音素境界ラベル1010を参照し、音声波形を音素単位1011に区切る。なお、合成単位作成部902は、指定された任意の合成単位に音声波形を区切ることができる。また、CV・VC単位1012とVCV単位1013の二つの合成単位を採用し、一つの音声波形辞書に複数の合成単位を持たせることも可能である。

【0115】候補波形抽出部903では、合成単位に区切られた音声波形を抽出し、この合成単位で抽出された音声波形の一覧を作成する。

【0116】もし、合成単位に複数の候補が存在する場合などは、登録波形選定部904で音声波形辞書に登録する音声波形を選択する。例として、図10の音声波形1000「家族の同意がある場合・・・(kazokuno/do-iga/arubaa)」を音素単位1011で区切った場合、この中に音素/k/の音声波形は、音声波形1021及び音声波形1022の2つ存在する。複数の音声波形を音声波形辞書に登録できる場合は別として、コンパクトな音声波形辞書を作成するためには、必要最低限の音声波形だけを登録することになる。したがって、登録波形選定部904では、候補となる複数の音声波形の中から、音声合成端末装置の利用者が設定した評価基準に基づき音声波形辞書に登録する音声波形を選択する。

【0117】登録波形選定部904で選択された音声波形は、パラメータ化部503によって音声合成器の条件に合致した所定の音声波形（パラメータ）形式に変換される。

【0118】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、音声合成端末装置の利用者に合成単位を自由に指定させることができ、例えば短い合成単位を指定されればサイズの小さな音声波形辞書を作成することができ、長い合成単位を指定されれば、高音質な音声合成が可能な音声波形辞書を作成することができる。したがって、本実施の形態に係る音声波形辞書作成装置によれば、音声合成端末装置の利用者は、音声波形辞書を記録するメモリ容量と、合成音声の音質との兼ね合いを判断して音声波形辞書を作成することが可能となる。

（第8の実施の形態）

【0119】まず、第8の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第7の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であるので、第7の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符号を付して詳細な説明を省略する。

【0120】図11に示すように、本実施の形態に係る音声波形辞書作成装置は、第7の実施の形態に係る音声波形辞書作成装置の登録波形選定部904（図9参照）の代わりに、登録波形選定部（登録波形選定手段）1101を備えている。

【0121】登録波形選定部1101は、候補波形抽出部903で抽出した各音声波形について、音素毎に長時間の平均スペクトル（スペクトル情報）を求めるスペクトル抽出部（スペクトル抽出手段）1102と、スペクトル抽出部1102で得た長時間平均スペクトルを用いて、同じ音素並びの音声波形間におけるスペクトル距離を計算するスペクトル距離計算部（スペクトル距離計算手段）1103と、スペクトル距離計算部1103で求めた音声波形間のスペクトル距離に基づき、音声波形辞書に代表として登録する音声波形を選定する代表波形選定部（スペクトル代表波形選定手段）1104とを備えている。

【0122】また、候補波形抽出部903は、音声データベース112a（図12参照）から音声波形を抽出し、この音声波形の中で音素並びが同じものが複数ある場合は、これらを音声波形辞書に登録する候補波形としてリストアップするようになっている。

【0123】また、パラメータ化部503は、登録波形選定部1101で選定された音声波形を、音声合成器の条件に合致した所定の波形（パラメータ）形式に変換するようになっている。

【0124】次に、本実施の形態に係る音声波形辞書作

成装置の動作について説明する。

【0125】候補波形抽出部903は、図12に示すように、音声データベース112aから音声合成端末装置の利用者が要求する合成単位で音声波形を抽出する。この際、音声波形に音素並びの同じものが複数ある場合は、これらの音声波形リストを作成する。例えば、候補波形抽出部903は、音声データベース112aに格納されている音声波形の中で必要な音素並びを含む音声波形の一覧、即ち、リスト1201に示す音声から、VCV連鎖を合成単位とする音声波形/ama/として、音声波形辞書に登録する候補波形である音声波形1211及び音声波形1212などの複数の音声波形を抽出する。

【0126】次いで、スペクトル抽出部1102は、候補波形抽出部903で抽出された複数の音声波形1211及び音声波形1212などの複数の音声波形について、各音素（/a/、/m/、/a/）毎にそのスペクトル情報1221～1226を計算する。なお、スペクトル情報は各音素内の長時間平均スペクトルである。

【0127】次いで、音声波形/ama/として複数存在する音声波形1211及び音声波形1212などの複数の音声波形の中から、音声波形辞書に登録する音声波形を選出する。そのためにスペクトル距離計算部1103では、候補音声波形間のスペクトル距離を評価関数として、候補音声波形をクラスタリングする。ここでクラスタリングによって生成されるクラスタ数は、音声波形辞書に登録する音声波形数であって、音声合成端末装置の利用者の要求する音声波形辞書のサイズによって決まる。

【0128】次いで、代表波形選定部1104では、スペクトル距離計算部1103によってクラスタリングされた各クラスタから代表音声波形1231及び代表音声波形1232を選出する。代表音声波形1231及び代表音声波形1232は、各クラスタにおける平均スペクトルに最も近い音声波形が選出される。

【0129】最後に、パラメータ化部503は、代表音声波形1231及び代表音声波形1232を音声波形辞書に登録する音声波形として所定の形式に変換する。

【0130】なお、上記説明では、VCV連鎖を合成単位とした場合を例にしたが、合成単位の種類にかかわらず本方式を用いることができる。

【0131】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、音声合成端末装置の利用者の要求する音声波形辞書のサイズに応じた音声波形数で、音声合成時に接続する二つの合成単位間で生じるスペクトル歪を少なくすることが可能な音声波形辞書を作成することができる。

（第9の実施の形態）

【0132】まず、第9の実施の形態に係る音声波形辞書作成装置の構成について説明する。なお、本実施の形態に係る音声波形辞書作成装置の構成は、第7の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様であ

るので、第7の実施の形態に係る音声波形辞書作成装置の構成とほぼ同様な構成については、同一の符合を付して詳細な説明を省略する。

【0133】図13に示すように、本実施の形態に係る音声波形辞書作成装置は、第7の実施の形態に係る音声波形辞書作成装置の登録波形選定部904（図9参照）の代わりに、登録波形選定部（登録波形選定手段）1301を備えている。

【0134】登録波形選定部1301は、候補波形抽出部903で抽出した各音声波形における基本周波数パターン（基本周波数情報）を求める基本周波数抽出部（基本周波数抽出手段）1302と、基本周波数抽出部1302で求めた音声波形の基本周波数パターンを考慮して、同じ音素並びの音声波形間の基本周波数距離の計算を行う基本周波数距離計算部（基本周波数距離計算手段）1303と、基本周波数距離計算部1303で求めた音声波形間の基本周波数距離に基づき、音声波形辞書に代表として登録する音声波形を選定する代表波形選定部（基本周波数代表波形選定手段）1304とを備えている。

【0135】また、候補波形抽出部903は、音声データベース112a（図9参照）から音声波形を抽出し、この音声波形の中で音素並びが同じものが複数ある場合は、これらを音声波形辞書に登録する候補波形としてリストアップするようになっている。

【0136】また、パラメータ化部503は、登録波形選定部1301で選定された音声波形を、音声合成器の条件に合致した所定の波形（パラメータ）形式に変換するようになっている。

【0137】次に、本実施の形態に係る音声波形辞書作成装置の動作について説明する。

【0138】まず、候補波形抽出部903は、音声データベース112aから、音声合成端末装置の利用者が要求する合成単位で音声波形を抽出する。この際、音声波形に音素並びの同じものが複数ある場合は、基本周波数パターンを考慮して音声波形辞書に登録する音声波形を選出することになる。なお、音声波形辞書に登録する音声波形数は、音声合成端末装置の利用者の要求する音声波形辞書サイズによって決まる。

【0139】次いで、基本周波数抽出部1302は、図14に示すように、候補波形抽出部903によって音素並び毎に抽出された音声波形について、その基本周波数パターン1401～1408を計算する。

【0140】次いで、基本周波数距離計算部1303は、基本周波数パターン1401～1408の基本周波数距離を類似度の評価関数として、音声波形をクラスタリングする。

【0141】次いで、代表波形選定部1304では、基本周波数距離計算部1303におけるクラスタリングによって生成される各クラス1421～1423から最も平均的な基本周波数パターンを代表音声波形1411～1

413として選出する。

【0142】最後に、パラメータ化部503は、代表音声波形1411～1413を音声波形辞書に登録する音声波形として所定の形式に変換する。

【0143】なお、上記説明では、VCV連鎖を合成単位とした場合を例にしたが、合成単位の種類にかかわらず本方式を用いることが可能である。

【0144】以上説明したように、本実施の形態に係る音声波形辞書作成装置は、音声合成端末装置の利用者の要求する音声波形辞書サイズに応じた音声波形数で、ピッチ変換や時間長変換などの韻律変換処理によって生じる波形歪を抑えることが可能な音声波形辞書を作成することができる。

（第10の実施の形態）

【0145】図15に示すように、第10の実施の形態に係る音声合成端末装置1501の構成は、第1の実施の形態に係る音声波形辞書配信システムの音声合成端末装置の構成と同様であるので、同一の符合を付して詳細な説明を省略する。

【0146】また、本実施の形態に係る音声合成端末装置1501は、第1の実施の形態に係る音声波形辞書配信システムの音声合成端末装置と同様に動作することができる。

【0147】したがって、音声合成端末装置1501は、利用者が音声波形辞書を決定することができ、情報提供サーバから獲得したテキスト情報を利用者の好みの話者の声で音声合成を行うことができる。また、音声合成端末装置1501を使用した音声波形辞書配信システムは、サーバ（音声波形辞書作成装置）側で音声合成して音声信号を送受信するシステムと比較して、音声波形辞書を一度ダウンロードすれば音声合成端末装置に音声信号を通信する必要があるため、通信負荷を抑えることができる。

（第11の実施の形態）

【0148】まず、第11の実施の形態に係る音声合成端末装置の構成について説明する。なお、本実施の形態に係る音声合成端末装置の構成は、第10の実施の形態に係る音声合成端末装置の構成とほぼ同様であるので、第10の実施の形態に係る音声合成端末装置の構成とほぼ同様な構成については、同一の符合を付して詳細な説明を省略する。

【0149】図16に示すように、本実施の形態に係る音声合成端末装置は、記録する音声波形辞書の数及びサイズを管理する音声波形辞書管理部1601を備えているとともに、第10の実施の形態に係る音声合成端末装置の記録媒体132（図15参照）の代わりに、音声波形辞書管理部1601によって管理された音声波形辞書を記録する記録媒体（音声波形辞書記録手段）1602を備えている。

【0150】記録媒体1602は、音声合成を行うため

に必要な音声波形情報によって構成された複数の音声波形辞書1603～1605を保存することが可能となっている。なお、音声波形辞書1603～1605の辞書形式は音声合成端末装置に搭載されている音声合成器に依存している。

【0151】次に、本実施の形態に係る音声合成端末装置の動作について説明する。

【0152】本実施の形態に係る音声合成端末装置は、音声波形辞書管理部1601で、記録媒体1602の記録容量をチェックすることによって記録媒体1602に記録可能な音声波形辞書の数及びサイズを管理し、利用者の要求によって音声波形辞書作成装置で作成した音声波形辞書をデータ送受信部135で受信し、記録媒体1602に格納する。

【0153】ここで、記録媒体1602では、上述したように、記録容量に応じて、一つまたは複数の音声波形辞書を格納することが可能である。

【0154】また、本実施の形態に係る音声合成端末装置は、音声合成部133において、記録媒体1602に記録してある音声波形辞書1603～1605の中から音声波形辞書を1つ選択し、選択した音声波形辞書を使用して音声合成を行う。

【0155】例えば、本実施の形態に係る音声合成端末装置は、音声読み上げの対象が電子メールの場合は話者1の音声波形辞書1603を使い、ニュース文の場合は話者2の音声波形辞書1604を使うなど、音声読み上げる情報毎に音声波形辞書を切り替えて使用することが可能である。

【0156】ただし、記録容量の限られた記録媒体1602で複数の音声波形辞書を記録するためには、音声波形辞書1つあたりのサイズを小さくする必要があるので、音質と、音声波形辞書の数とはトレードオフの関係になる。

【0157】以上説明したように、本実施の形態に係る音声合成端末装置は、記録媒体1602の記録容量と、記録媒体1602に記録する音声波形辞書のサイズとに応じて、記録媒体1602に記録する音声波形辞書の数を決めることができるので、複数の音声波形辞書を保持することができる。したがって、本実施の形態に係る音声合成端末装置は、一つのアプリケーションの中で複数話者での合成が可能となり、例えば電子メールやニュース文などコンテンツ毎に異なる話者での音声読み上げが可能となるので、利用者は現在の音声読み上げの対象が何であるのかを判断することができる。

(第12の実施の形態)

【0158】まず、第12の実施の形態に係る音声合成端末装置の構成について説明する。なお、本実施の形態に係る音声合成端末装置の構成は、第11の実施の形態に係る音声合成端末装置の構成とほぼ同様であるので、第11の実施の形態に係る音声合成端末装置の構成とほ

ぼ同様な構成については、同一の符号を付して詳細な説明を省略する。

【0159】図17に示すように、本実施の形態に係る音声合成端末装置は、第11の実施の形態に係る音声合成端末装置の音声合成部133（図16参照）の代わりに、LPC方式を代表とするパラメトリックな音声合成器や波形重畳系のノンパラメトリックな音声合成器など、複数の音声合成器によって構成される音声合成部（音声合成手段）1701を備えている。

【0160】例えば、音声合成部1701は、音声合成器1702～1704によって構成されており、音声合成器1702～1704は、それぞれが採用している音声合成方式に適したフォーマットの音声波形辞書を参照して音声合成を行うようになっている。

【0161】なお、音声波形辞書管理部1601は、複数の音声合成器1702～1704の中から、記録媒体1602に記録された音声波形辞書1603及び1604などの複数の音声波形辞書のうち、音声合成に使用する音声波形辞書の辞書形式に適合した音声合成器を選択し、選択した音声合成器と、音声合成に使用する音声波形辞書との間で音声波形データの受け渡しをするようになっている。

【0162】次に、本実施の形態に係る音声合成端末装置の動作について説明する。

【0163】本実施の形態に係る音声合成端末装置は、例えば、音声読み上げの対象が電子メールの場合は音声波形辞書（話者1）1603を利用し、ニュース文の場合は音声波形辞書（話者2）1604を利用するというように、利用者の設定によって音声読み上げの対象コンテンツ毎にそれぞれの音声波形辞書が割り振られている場合、音声読み上げの対象コンテンツ毎に使用する音声波形辞書を音声波形辞書管理部1601によって切り替える。

【0164】ここで、音声読み上げに使用する音声波形辞書（話者1）1603がピッチ同期の波形重畳型合成方式に対応した辞書形式であった場合、音声波形辞書管理部1601によって、この合成方式を扱える音声合成器1702が選択され、音声波形辞書（話者2）1604がLSP係数を利用したフレーム同期型の合成方式に対応した辞書形式であった場合、音声波形辞書管理部1601によって、この合成方式を扱える音声合成器1703が選択される。

【0165】このように、音声合成部1701は、音声波形辞書管理部1601によって、音声読み上げに使用される音声波形辞書の辞書形式に応じて音声合成器が自動的に選択され、選択された音声合成器によって音声合成を実現する。

【0166】したがって、音声波形辞書1603及び音声波形辞書1604の辞書形式としては、音声合成端末装置に搭載されている音声合成器に対応した辞書形式の

みが許される。

【0167】以上説明したように、本実施の形態に係る音声合成端末装置は、例えば音声波形辞書の辞書形式がLPCパラメータである場合や、音声波形をピッチ波形単位に分割した形式である場合など、それぞれの辞書形式に応じて音声合成器を自動的に使い分けることで、音声波形辞書をダウンロードする際、利用者に辞書形式を把握させる必要がない。

【0168】なお、本実施の形態においては、音声合成端末装置に波形重畳型の合成方式に対応した音声合成器や、LSP係数を利用したフレーム同期型の合成方式に対応した音声合成器を搭載した例を示したが、本発明によれば、音声合成端末装置には、これらの音声合成器以外にも様々な音声合成方式に対応した音声合成器を搭載することが可能である。

【0169】また、本実施の形態においては、記録媒体1602が複数の音声波形辞書を記録する構成を採っていたが、本発明によれば、記録媒体1602は音声波形辞書を1つだけ記録する構成であっても良い。

(第13の実施の形態)

【0170】まず、第13の実施の形態に係る音声合成端末装置の構成について説明する。なお、本実施の形態に係る音声合成端末装置の構成は、第10の実施の形態に係る音声合成端末装置の構成とほぼ同様であるので、第10の実施の形態に係る音声合成端末装置の構成とほぼ同様な構成については、同一の符号を付して詳細な説明を省略する。

【0171】図18に示すように、本実施の形態に係る音声合成端末装置は、第10の実施の形態に係る音声合成端末装置の音声合成部133(図15参照)の代わりに、音声合成部(音声合成手段)1801を備えている。

【0172】音声合成部1801は、サーバなどから配信されたテキスト情報を、形態素解析や構文解析などを行って、韻律情報を付属させた表音文字並びに変換するテキスト解析部(テキスト解析手段)1802と、テキスト解析部1802によるテキスト解析の結果から、音声合成対象となる表音文字並び相応しい韻律パターン(基本周波数パターンと音韻継続時間長パターン、振幅パターン)を生成する韻律生成部(韻律生成手段)1803と、指定された合成単位毎に表音文字並びが一致する音声波形を後述する音声波形辞書1814の中から検索する音声波形選択部(音声波形選択手段)1804と、音声波形選択部1804で選択された音声波形について、ターゲットとなる韻律パターンに沿って基本周波数や音韻継続時間長の変更を行い、最終的に音声波形を接続して合成音声出力する音声波形変形部(音声波形変形手段及び音声波形接続手段)1805とを備えている。

【0173】なお、音声波形選択部1804は、複数の音声波形候補が存在する場合は、合成単位長や韻律パタ

ンなどの条件が最も適合する音声波形を選択して用いるようになっている。

【0174】また、本実施の形態に係る音声合成端末装置は、第10の実施の形態に係る音声合成端末装置の記録媒体132(図15参照)の代わりに、記録媒体(音声波形辞書記録手段)1811を備えている。

【0175】記録媒体1811は、テキスト解析部1802で用いる単語辞書であり、読み情報、品詞情報、アクセント情報などを保持するテキスト解析辞書1812と、韻律生成部1803が基本周波数パターンや音韻継続時間長パターンを決定するのに必要なパラメータを保持する韻律辞書1813と、音声合成部1801が参照可能な所定の形式の音声波形データによって構成される音声波形辞書1814とを備えている。

【0176】次に、本実施の形態に係る音声合成端末装置の動作について説明する。

【0177】本実施の形態に係る音声合成端末装置は、情報提供サーバなどから配信されたテキスト情報をデータ送受信部135で受信すると、受信したテキスト情報を音声合成部1801に出力する。音声合成部1801では、データ送受信部135から出力されたテキスト情報を以下のように処理する。

【0178】まず、テキスト解析部1802が、形態素解析を行って、単語辞書(テキスト解析辞書1812)を参照して表音文字並びを生成した後、構文解析を行って、生成した表音文字並びにアクセントやポーズなどの韻律情報を付属させる。

【0179】次いで、韻律生成部1803が、アクセントフレーズ毎の基本周波数パターンや音素又は音韻毎の継続時間長テーブルを保持する韻律辞書1813を参照することで、合成する表音文字並びに適した基本周波数パターンや音韻継続時間長パターンを生成する。

【0180】次いで、音声波形選択部1804が、合成する表音文字並びを合成単位に分割し、これと一致する音声波形を音声波形辞書1814の中から検索する。もし、一つの表音文字並びに対して複数の音声波形候補が存在する場合は、韻律パターンの適合性や接続位置でのスペクトルの適合性を評価関数として、合成環境に最も適合した音声波形を選択する。

【0181】次いで、音声波形変形部1805が、音声波形選択部1804によって選択された音声波形の韻律をターゲットの韻律パターンに変換する。例として、音声合成部1801がピッチ同期の波形重畳型音声合成方式を採用している場合、音声波形変形部1805では必要なピッチ波形を音声波形辞書1814から読み出し、このピッチ波形をターゲットの基本周波数間隔に合わせて配列することで基本周波数変更を行う。また、ピッチ波形を所望する時間長だけ配列することで音韻継続時間長を制御する。

【0182】最後に、音声合成端末装置は、韻律変換し

た音声波形を音声波形変形部1805で接続してスピーカ134から出力することによって、情報提供サーバなどから配信されたテキスト情報を音声読み上げる。

【0183】以上説明したように、本実施の形態に係る音声合成端末装置は、音声波形辞書に登録されている複数の音声波形の候補の中から、合成環境に最も適した音声波形を選択することができるようになり、高音質な音声合成が可能となる。

【0184】なお、上記説明では音声波形変形部1805に波形重畳型の音声合成方式を採用した場合を例にしたが、その他パラメトリックな音声合成方式を採用することも可能である。

【0185】

【発明の効果】以上説明したように、本発明によれば、利用者の要求する話者の音色で電子メールなどのテキスト情報を読み上げることが可能な音声波形辞書配信システム、音声波形辞書作成装置、及び音声合成端末装置を提供することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態に係る音声波形辞書配信システムのブロック図

【図2】本発明の第2の実施の形態に係る音声波形辞書配信システムのブロック図

【図3】図2に示す音声波形辞書配信システムの動作のフローチャート

【図4】本発明の第3の実施の形態に係る音声波形辞書作成装置のブロック図

【図5】本発明の第4の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図6】本発明の第5の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図7】本発明の第6の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図8】図7に示す音声波形辞書作成装置の動作を説明する模式図

【図9】本発明の第7の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図10】図9に示す音声波形辞書作成装置の動作を説明する模式図

【図11】本発明の第8の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図12】図11に示す音声波形辞書作成装置の動作を説明する模式図

【図13】本発明の第9の実施の形態に係る音声波形辞書作成装置の一部のブロック図

【図14】図13に示す音声波形辞書作成装置の動作を説明する模式図

【図15】本発明の第10の実施の形態に係る音声合成端末装置のブロック図

【図16】本発明の第11の実施の形態に係る音声合成

端末装置の一部のブロック図

【図17】本発明の第12の実施の形態に係る音声合成端末装置の一部のブロック図

【図18】本発明の第13の実施の形態に係る音声合成端末装置の一部のブロック図

【符号の説明】

100	音声波形辞書配信システム
110	音声波形辞書作成装置
111	辞書条件決定部（辞書条件決定手段）
112 a、112 b、112 c	音声データベース
113	音声波形辞書作成部（音声波形辞書作成手段）
114	データ送受信部（条件受信辞書送信手段）
120	ネットワーク
130、140、150	音声合成端末装置
131	辞書条件入力部（辞書条件入力手段）
132	記録媒体（音声波形辞書記録手段）
133	音声合成部（音声合成手段）
134	スピーカ（音声出力手段）
135	データ送受信部（条件送信辞書受信手段）
200	音声波形辞書配信システム
201	音声波形辞書作成装置
202	音声合成部（作成辞書音声合成手段）
401	音声波形辞書作成装置
501	音声波形辞書作成部（音声波形辞書作成手段）
502	音声波形選出部（音声波形選出手段）
503	パラメータ化部（パラメータ化手段）
504	音声波形付属情報抽出部（音声波形付属情報抽出手段）
505	辞書構築部（辞書構築手段）
601	パラメータ化部（パラメータ化手段）
611	音声波形
612	スペクトル包絡（声道特性）
613	音源パルス列（音源特性）
701	パラメータ化部（パラメータ化手段）
702	ピッチ波形抽出部（ピッチ波形抽出手段）
703	ピッチ波形分類部（ピッチ波形分類手段）
704	代表ピッチ波形作成部（代表ピッチ波形作成手段）
801	音声波形
821、822	代表ピッチ波形
901	音声波形選出部（音声波形選出手段）
902	合成単位作成部（合成単位作成手段）
903	候補波形抽出部（候補波形抽出手段）
904	登録波形選定部（登録波形選定手段）
1000	音声波形
1010	音素境界ラベル
1011	音素単位（合成単位）
1012	C V・V C単位（合成単位）

31

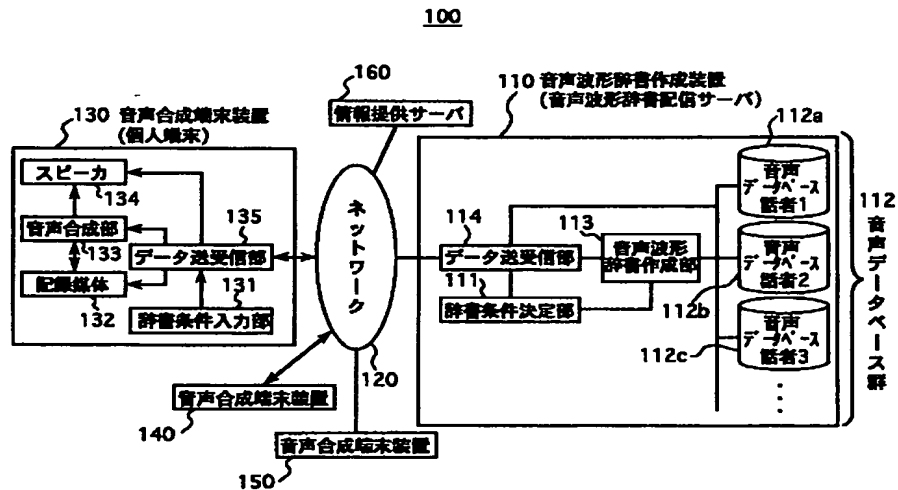
32

1013 VCV単位(合成単位)
 1021、1022 音声波形
 1101 登録波形選定部(登録波形選定手段)
 1102 スペクトル抽出部(スペクトル抽出手段)
 1103 スペクトル距離計算部(スペクトル距離計算手段)
 1104 代表波形選定部(スペクトル代表波形選定手段)
 1211、1212 音声波形
 1221 スペクトル情報
 1301 登録波形選定部(登録波形選定手段)
 1302 基本周波数抽出部(基本周波数抽出手段)
 1303 基本周波数距離計算部(基本周波数距離計算手段)

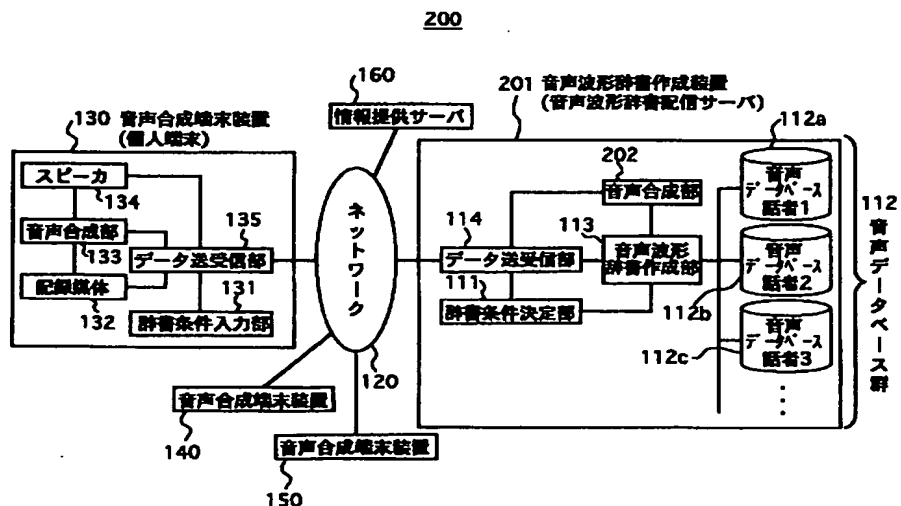
*1304 代表波形選定部(基本周波数代表波形選定手段)
 1401 基本周波数パターン(基本周波数情報)
 1501 音声合成端末装置
 1602 記録媒体(音声波形辞書記録手段)
 1603、1604、1605 音声波形辞書
 1701 音声合成部(音声合成手段)
 1702、1703、1704 音声合成器
 1801 音声合成部(音声合成手段)
 1802 テキスト解析部(テキスト解析手段)
 1803 韻律生成部(韻律生成手段)
 1804 音声波形選択部(音声波形選択手段)
 1805 音声波形変形部(音声波形変形手段、音声波形接続手段)
 1811 記録媒体(音声波形辞書記録手段)

*

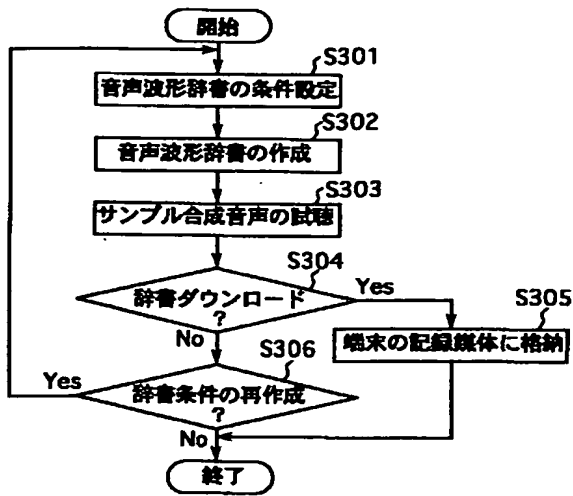
【図1】



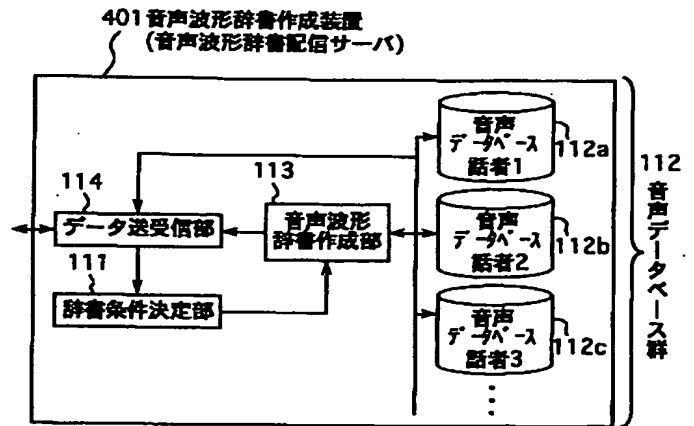
【図2】



【図3】

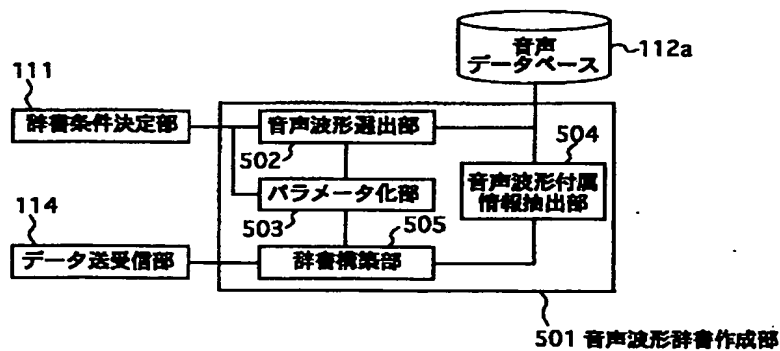


【図4】

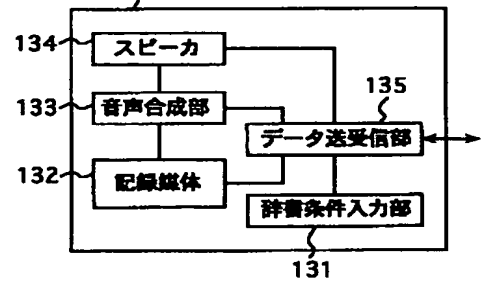


【図15】

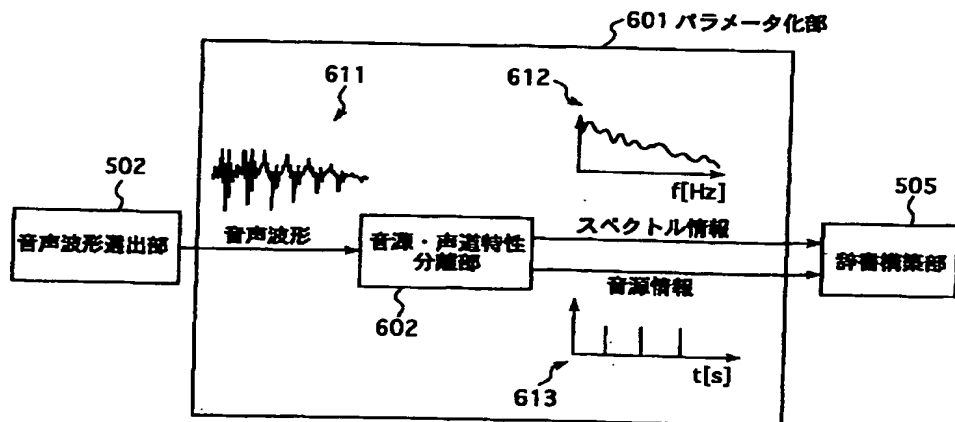
【図5】



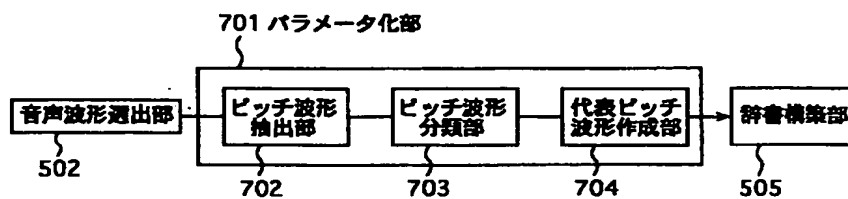
1501 音声合成端末装置(個人端末)



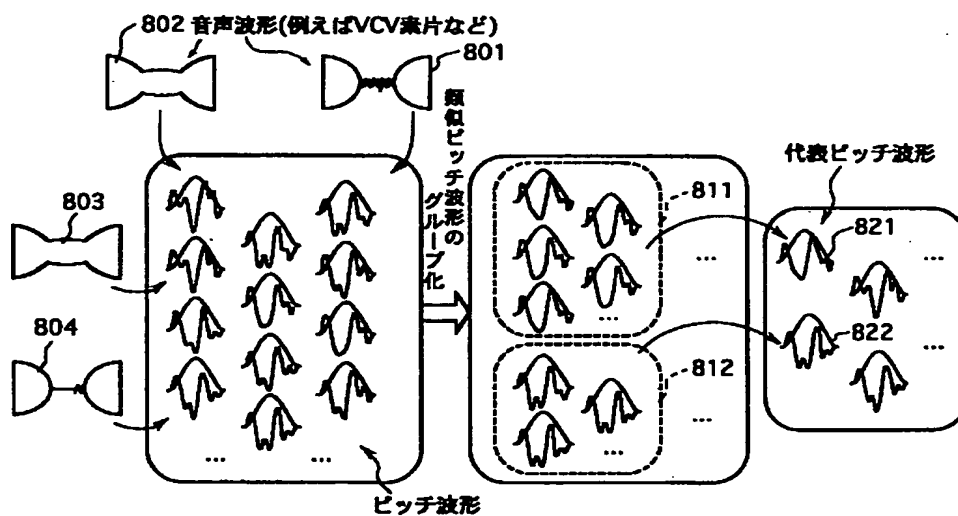
【図6】



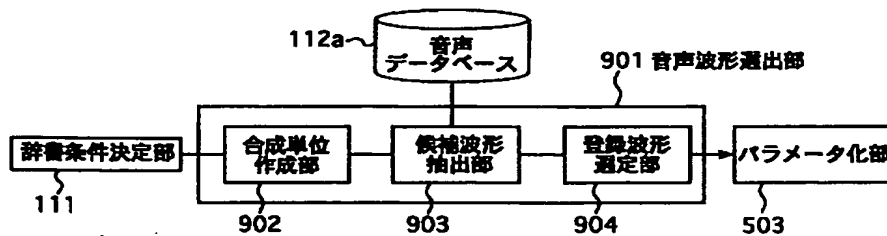
【図7】



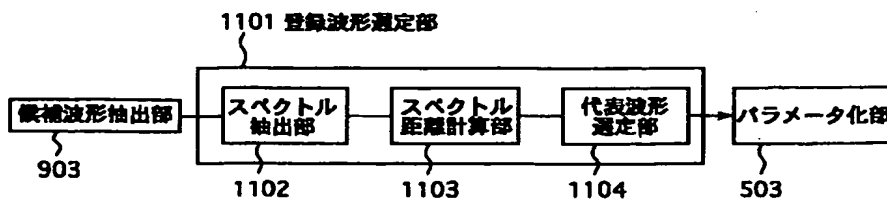
【図8】



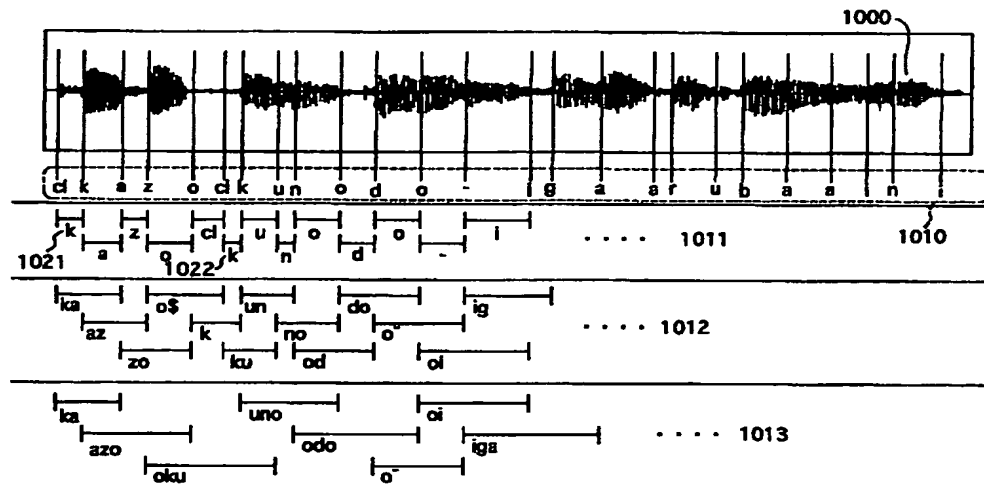
【図9】



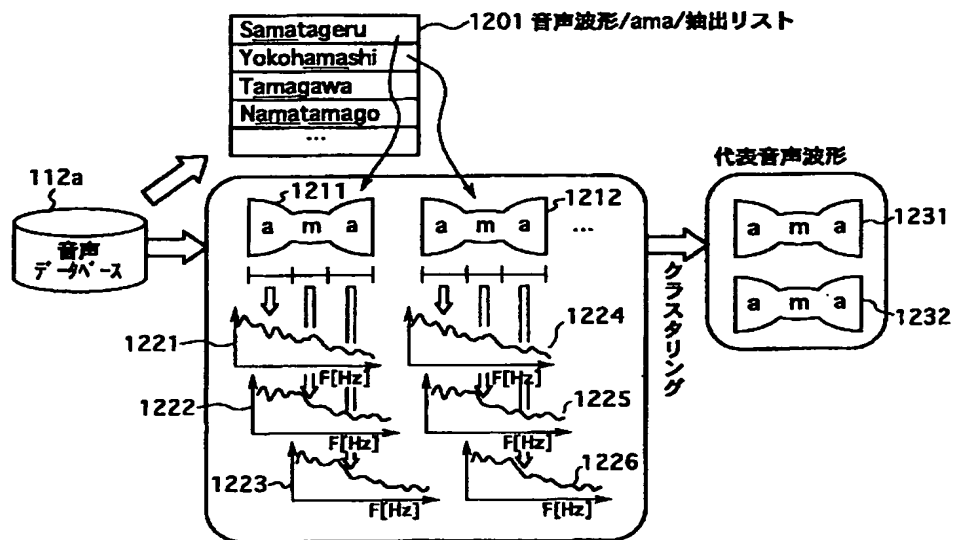
【図11】



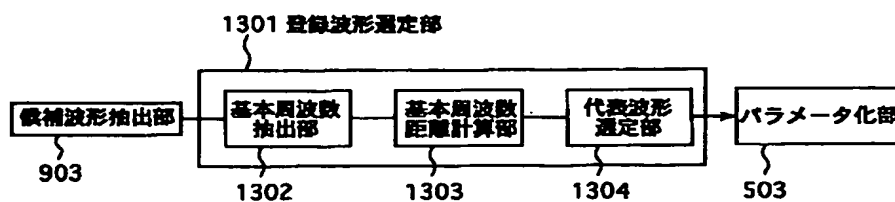
【図10】



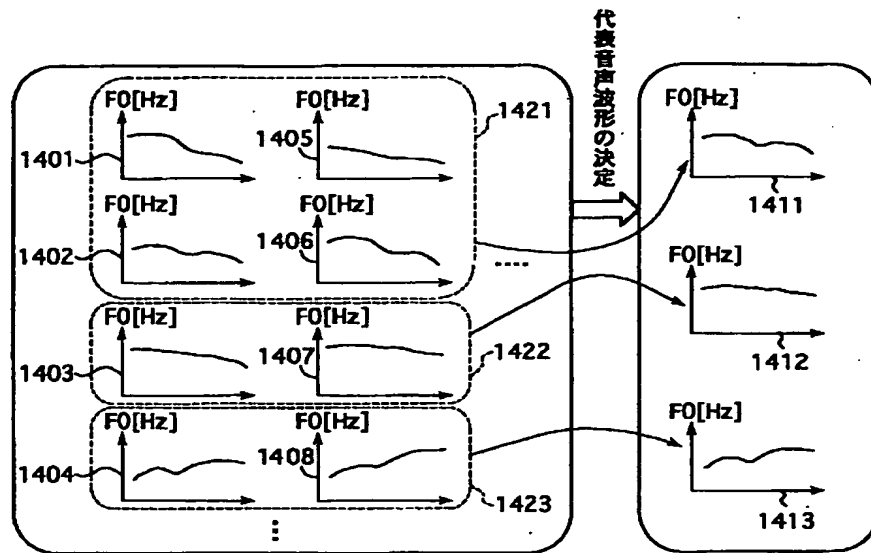
【図12】



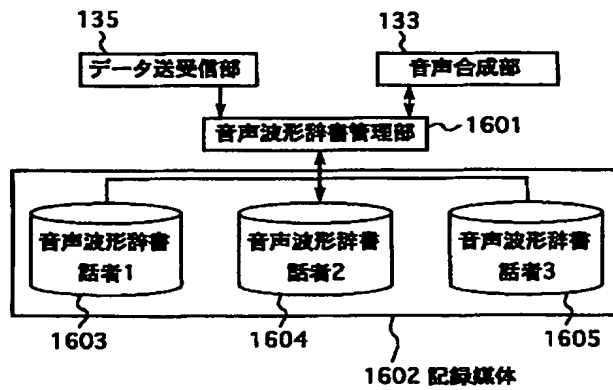
【図13】



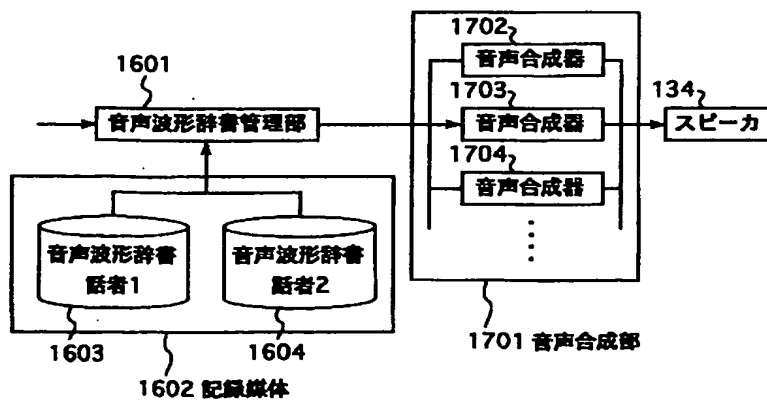
【図14】



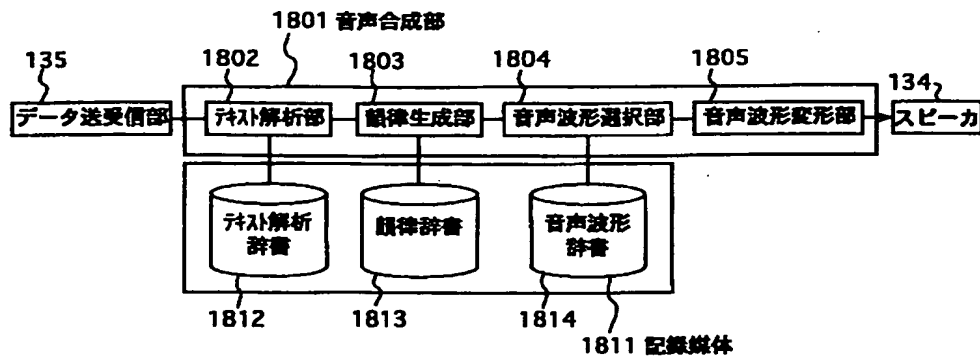
【図16】



【図17】



【図18】



フロントページの続き

(51)Int.Cl.⁷

識別記号

F I
G I O L 3/00

テーマコード(参考)
N